```
   1 GGCGTCGCCG CGCGGGGAGA AAGAAGCCGC GCCCAGCCCG GCGTCCCGAG
  51 CAGCGCAGGG GAGGATCCCC GCGCAGTGAC CCGGGAGCCA CCACAGACTC
 101 TGGGAGGCTC GGCGGCTGGA GCAGCAGGCA GCTCCCCGCA GCTCCCGGCG
 151 CTTCCAGGCA GCTCTCTGAG CCGTGCCAGA GGCCCGGCCC GCCATTCCCA
 201 GCCCCGAGCC ATGATGAAGA CTTTGTCCAG CGGGAACTGC ACGCTCAGTG
 251 TGCCCGCCAA AAACTCATAC CGCATGGTGG TGCTGGGTGC CTCTCGGGTG
 301 GGCAAGAGCT CCATCGTGTC TCGCTTCCTC AATGGCCGCT TTGAGGACCA
 351 GTACACACCC ACCATCGAGG ACTTCCACCG TAAGGTATAC AACATCCGCG
 401 GCGACATGTA CCAGCTCGAC ATCCTGGATA CCTCTGGCAA CCACCCCTTC
 451 CCCGCCATGC GCAGGCTGTC CATCCTCACA GGGGATGTCT TCATCCTGGT
 501 GTTCAGCCTG GATAACCGGG AGTCCTTCGA TGAGGTCAAG CGCCTTCAGA
 551 AGCAGATCCT GGAGGTCAAG TCCTGCCTGA AGAACAAGAC CAAGGAGGCG
 601 GCGGAGCTGC CCATGGTCAT CTGTGGCAAC AAGAACGACC ACGGCGAGCT
 651 GTGCCGCCAG GTGCCCACCA CCGAGGCCGA GCTGCTGGTG TCGGGCGACG
 701 AGAACTCCGC CTACTTCGAG GTGTCGGCCA AGAAGAACAC CAACGTGGAC
 751 GAGATGTTCT ACGTGCTCTT CAGCATGGCC AAGCTGCCAC ACGAGATGAG
 801 CCCCGCCCTG CATCGCAAGA TCTCCGTGCA GTACGGTGAC GCCTTCCACC
 851 CCAGGCCCTT CTGCATGCGC CGCGTCAAGG AGATGGACGC CTATGGCATG
 901 GTCTCGCCCT TCGCCCGCCG CCCCAGCGTC AACAGTGACC TCAAGTACAT
 951 CAAGGCCAAG GTCCTTCGGG AAGGCCAGGC CCGTGAGAGG GACAAGTGCA
1001 CCATCCAGTG AGCGAGGGAT GCTGGGGCGG GGCTTGGCCA GTGCCTTCAG
1051 GGAGGTGGCC CCAGATGCCC ACTGTGCGCA TCTCCCCACC GAGGCCCCGG
1101 CAGCAGTCTT GTTCACAGAC CTTAGGCACC AGACTGGAGG CCCCCGGGCG
1151 CTGGCCTCCG CACATTCGTC TGCCTTCTCA CAGCTTTCCT GAGTCCGCTT
1201 GTCCACAGCT CCTTGGTGGT TTCATCTCCT CTGTGGGAGG ACACATCTCT
1251 GCAGCCTCAA GAGTTAGGCA GAGACTCAAG TTACACCTTC CTCTCCTGGG
1301 GTTGGAAGAA ATGTTGATGC CAGAGGGGTG AGGATTGCTG CGTCATATGG
1351 AGCCTCCTGG GACAAGCCTC AGGATGAAAA GGACACAGAA GGCCAGATGA
1401 GAAAGGTCTC CTCTCTCCTG GCATAACACC CAGCTTGGTT TGGGTGGCAG
1451 CTGGGAGAAC TTCTCTCCCA GCCCTGCAAC TCTTACGCTC TGGTTCAGCT
1501 GCCTCTGCAC CCCCTCCCAC CCCCAGCACA CACACAAGTT GGCCCCCAGC
1551 TGCGCCTGAC ATTGAGCCAG TGGACTCTGT GTCTGAAGGG GGCGTGGCCA
1601 CACCTCCTAG ACCACGCCCA CCACTTAGAC CACGCCCACC TCCTGACCGC
1651 GTTCCTCAGC CTCCTCTCCT AGGTCCCTCC GCCCGACAGT TGTGCTTTGT
1701 TGTGGTTGCA GCTGTTTTCG TGTCATGTAT AGTAGTAGAA ATGGAAATCA
1751 TTGTACTGTA AAAGCCTAGT GACTCCCTCC TTGGCCAGGC CCTCACCCAG
1801 TTCAGATCCA CGGCCTCCAC CCGGGACGCC TTCCTCCTCT GCTCCCAAAC
1851 AGGGTTTCCG TGGCCTGTTT GCAGCTAGAC ATTGACCTCC GCCATTGAGC
1901 TCCACGGTTT ACAGACAATT GCACAAGCGT GGGGTGGGCA GGCCAGGACT
1951 GCTTTTTTTT AATGCTCCCA TTTCACAGAG GATACCACCG AGACTCGGAG
2001 GGGACACGAT GAGCACCAGG CCCCACCTTT GTCCCCTAGC AAATTCAGGG
2051 TACAGCTCCA CCTAGAACCA GGCTGCCCTC TACTGTGCTC GTTCCTCAAG
2101 CATTTATTAA GCACCTACTG GGTGCTGGGT TCACTGTGTC CTAGGAAACC
2151 AAGAGGGTCC CCAGTCCTGG CCTCTGCCCG CCCCTGCTGC CCCACCACCT
2201 TCTGCACACA CAGCGGTGGG GAGGCGGGGA GGAGCAGCTG GGACCCAGAA
2251 CTGAGCCTGG GAGGGATCCG ACAGAAAAGC TCAGGGCGGG TCTTCTCCTT
2301 GTGCCCGGGA TTGGGCTATG CTGGGTACCA CCATGTACTC AGGCATGGTG
2351 GGTTTTGAAC CCATAAACCA AAGGCCCTTG TCATCAGCTC TTAACAAGTA
2401 TATTTTGTAT TTTAATCTCT CTAAACATAT TGAAGTTTTA GGGCCCTAAG
2451 GAACCTTAGT GATCTTCTAT TGGGTCTTTC TGAGGTTCAG AGAGGGTAAG
2501 TAACTTCCTC CAGGTCACAC AGCAAGTCTG TGGGTGGCAG AAGCAAGCTA
2551 GCGCTGGGCA TTCAGTACAT ACCACGATGT GCTCCCTCTC TTGATGCTTG
2601 GCCCCTGGGG CCTTCAGGGC TTTGGGACAT CTTGTCCTCA ACCCTCTCCC
2651 TAGATCAGTC TGTGAGGGTC CCTGTAGATA TTGTGTACAC CATGCCCATG
2701 TATATACAAG TACACACAGA TGTACACACA GATGTACACA TGCTCCAGCC
2751 CCAGCTCTGC ATACCTGCAC CTGCACCCCA GCCTTGGCCC CTGCCTGCGT
2801 CTGTGCTCAA AGCAGCAGCT CCAACCCTGC CTCTGTCCCC TTCCCACCC
2851 ACTGCCTGAG CCTTCTGAGC AGACCAGGTA CCTTGGCTGC ACCGGTGTGT
2901 GGCCCGCTCT CACCCAGGCA CAGCCCCGCC ACCATGGATC TCCGTGTACA
2951 CTATCAATAA AAGTGGGTTT GTTACAAAAA AAAAAAAAA AAAAAAAAAA
3001 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA
3051 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AA
```

FEATURES:
5'UTR: 1-210
Start Codon: 211
Stop Codon: 1009
3'UTR: 1012

5' UTR ANALYSIS:
Query=cDNA clone
Sbjct: genomic sequence

5' UTR Exon 1, non-coding
Score = 58.0 bits (29), Expect = 2e-10
 Identities = 29/29 (100%)
 Strand = Plus / Plus

Query: 1   ggcgtcgccgcgcggggagaaagaagccg 29
             |||||||||||||||||||||||||||||
Sbjct: 1535 ggcgtcgccgcgcggggagaaagaagccg 1563

5' UTR Exon 2, non-coding
  Score = 349 bits (176), Expect = 3e-98
 Identities = 176/176 (100%)
 Strand = Plus / Plus

Query: 26   gccgcgcccagcccggcgtcccgagcagcgcaggggaggatccccgcgcagtgacccggg 85
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 2001 gccgcgcccagcccggcgtcccgagcagcgcaggggaggatccccgcgcagtgacccggg 2060

Query: 86   agccaccacagactctgggaggctcggcggctggagcagcaggcagctccccgcagctcc 145
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 2061 agccaccacagactctgggaggctcggcggctggagcagcaggcagctccccgcagctcc 2120

Query: 146   cggcgcttccaggcagctctctgagccgtgccagaggcccggcccgccattcccag 201
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 2121 cggcgcttccaggcagctctctgagccgtgccagaggcccggcccgccattcccag 2176

  Score = 563 bits (284), Expect = e-162
 Identities = 284/284 (100%)
 Strand = Plus / Plus

5' UTR Exon 3, Protein coding region begins at nucleotide 211

Query: 199   cagccccgagccatgatgaagactttgtccagcgggaactgcacgctcagtgtgcccgcc 258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 7474 cagccccgagccatgatgaagactttgtccagcgggaactgcacgctcagtgtgcccgcc 7533

Homologous proteins:
Top BLAST Hits

	Score	E
CRA|18000005194969 /altid=gi|10047088 /def=ref|NP_055125.1| sim...	538	e-152
CRA|18000005238449 /altid=gi|5059122 /def=gb|AAD38928.1|AF13440...	512	e-144
CRA|332000009620725 /altid=gi|8118457 /def=gb|AAF72997.1|AF2620...	342	2e-93
CRA|18000005232775 /altid=gi|7706359 /def=ref|NP_057168.1| ras-...	342	2e-93
CRA|87000000006130 /altid=gi|7230768 /def=gb|AAF43090.1|AF23915...	342	3e-93
CRA|18000005090459 /altid=gi|6677673 /def=ref|NP_033052.1| RAS,...	341	6e-93
CRA|89000000197633 /altid=gi|7295299 /def=gb|AAF50620.1| (AE003...	228	8e-59
CRA|105000014645240 /altid=gi|10503969 /def=gb|AAG17979.1|AF177...	214	1e-54

```
BLAST dbEST hits:
gi|9345313 /dataset=dbest /taxon=960...                        880  0.0
gi|9335874 /dataset=dbest /taxon=960...                        846  0.0
gi|10143211 /dataset=dbest /taxon=96...                        821  0.0
gi|9335309 /dataset=dbest /taxon=960...                        813  0.0
gi|9150610 /dataset=dbest /taxon=9606...                       662  0.0
gi|10144589 /dataset=dbest /taxon=96...                        617  e-174
gi|9333908 /dataset=dbest /taxon=960...                        599  e-169
```

EXPRESSION INFORMATION FOR MODULATORY USE:

library source:

Expression information from BLAST dbEST hits:

```
gi|9345313   Placenta choriocarcinoma
gi|9335874   Uterus endometrium adenocarcinoma
gi|10143211  Skin melanotic melanoma
gi|9335309   Uterus endometrium adenocarcinoma
gi|9150610   Skin melanotic melanoma
gi|10144589  Skin melanotic melanoma
gi|9333908   Uterus endometrium adenocarcinoma
```

Expression information from PCR-based tissue screening panels:

Human fetal whole brain

FIGURE 1, page 3 of 3

```
  1 MMKTLSSGNC TLSVPAKNSY RMVVLGASRV GKSSIVSRFL NGRFEDQYTP
 51 TIEDFHRKVY NIRGDMYQLD ILDTSGNHPF PAMRRLSILT GDVFILVFSL
101 DNRESFDEVK RLQKQILEVK SCLKNKTKEA AELPMVICGN KNDHGELCRQ
151 VPTTEAELLV SGDENSAYFE VSAKKNTNVD EMFYVLFSMA KLPHEMSPAL
201 HRKISVQYGD AFHPRPFCMR RVKEMDAYGM VSPFARRPSV NSDLKYIKAK
251 VLREGQARER DKCTIQ
```

FEATURES:
Functional domains and key regions:
[1] PDOC00001 PS00001 ASN_GLYCOSYLATION
N-glycosylation site

Number of matches: 2
```
    1        9-12 NCTL
    2      125-128 NKTK
```

[2] PDOC00004 PS00004 CAMP_PHOSPHO_SITE
cAMP- and cGMP-dependent protein kinase phosphorylation site

Number of matches: 4
```
    1       84-87 RRLS
    2      174-177 KKNT
    3      202-205 RKIS
    4      236-239 RRPS
```

[3] PDOC00005 PS00005 PKC_PHOSPHO_SITE
Protein kinase C phosphorylation site

Number of matches: 2
```
    1       19-21 SYR
    2      172-174 SAK
```

[4] PDOC00006 PS00006 CK2_PHOSPHO_SITE
Casein kinase II phosphorylation site

Number of matches: 5
```
    1       51-54 TIED
    2      105-108 SFDE
    3      154-157 TEAE
    4      161-164 SGDE
    5      177-180 TNVD
```

[5] PDOC00007 PS00007 TYR_PHOSPHO_SITE
Tyrosine kinase phosphorylation site

```
        220-228 RRVKEMDAY
```

[6] PDOC00266 PS00294 PRENYLATION
Prenyl group binding site (CAAX box)

```
        263-266 CTIQ
```

[7] PDOC00016 PS00016 RGD
Cell attachment sequence

```
         63-65 RGD
```

[8] PDOC00017 PS00017 ATP_GTP_A
ATP/GTP-binding site motif A (P-loop)

```
         26-33 GASRVGKS
```

FIGURE 2, page 1 of 3

BLAST Alignment to Top Hit:
>CRA|18000005194969 /altid=gi|10047088 /def=ref|NP_055125.1| similar
          to mouse Ras, dexamethasone-induced 1; tumor endothelial
          marker 2 [Homo sapiens] /org=Homo sapiens /taxon=9606
          /dataset=nraa /length=278
          Length = 278

 Score = 538 bits (1372), Expect = e-152
 Identities = 265/266 (99%), Positives = 265/266 (99%)

Query: 1   MMKTLSSGNCTLSVPAKNSYRMVVLGASRVGKSSIVSRFLNGRFEDQYTPTIEDFHRKVY 60
            MMKTLSSGNCTLSVPAKNSYRMVVLGASRVGKSSIVSRFLNGRFEDQYTPTIEDFHRKVY
Sbjct: 13   MMKTLSSGNCTLSVPAKNSYRMVVLGASRVGKSSIVSRFLNGRFEDQYTPTIEDFHRKVY 72

Query: 61   NIRGDMYQLDILDTSGNHPFPAMRRLSILTGDVFILVFSLDNRESFDEVKRLQKQILEVK 120
            NIRGDMYQLDILDTSGNHPFPAMRRLSILTGDVFILVFSLDNRESFDEVKRLQKQILEVK
Sbjct: 73   NIRGDMYQLDILDTSGNHPFPAMRRLSILTGDVFILVFSLDNRESFDEVKRLQKQILEVK 132

Query: 121 SCLKNKTKEAAELPMVICGNKNDHGELCRQVPTTEAELLVSGDENSAYFEVSAKKNTNVD 180
            SCLKNKTKEAAELPMVICGNKNDHGELCRQVPTTEAELLVSGDEN AYFEVSAKKNTNVD
Sbjct: 133 SCLKNKTKEAAELPMVICGNKNDHGELCRQVPTTEAELLVSGDENCAYFEVSAKKNTNVD 192

Query: 181 EMFYVLFSMAKLPHEMSPALHRKISVQYGDAFHPRPFCMRRVKEMDAYGMVSPFARRPSV 240
            EMFYVLFSMAKLPHEMSPALHRKISVQYGDAFHPRPFCMRRVKEMDAYGMVSPFARRPSV
Sbjct: 193 EMFYVLFSMAKLPHEMSPALHRKISVQYGDAFHPRPFCMRRVKEMDAYGMVSPFARRPSV 252

Query: 241 NSDLKYIKAKVLREGQARERDKCTIQ 266
            NSDLKYIKAKVLREGQARERDKCTIQ
Sbjct: 253 NSDLKYIKAKVLREGQARERDKCTIQ 278

Complete Amino Acid Sequence of gi|10047088 /def=ref|NP_055125.1| similar
          to mouse Ras, dexamethasone-induced 1; tumor endothelial
          marker 2 [Homo sapiens] /org=Homo sapiens /taxon=9606
          /dataset=nraa /length=278
          Length = 278
NOTE: UNDERLINED RESIDUES NOT PRESENT IN THE PROTEIN OF THE PRESENT INVENTION
        1 mpaslallqp rammktlssg nctlsvpakn syrmvvlgas rvgkssivsr flngrfedqy
       61 tptiedfhrk vynirgdmyq ldildtsgnh pfpamrrlsi ltgdvfilvf sldnresfde
      121 vkrlqkqile vksclknktk eaaelpmvic gnkndhgelc rqvptteael lvsgdencay
      181 fevsakkntn vdemfyvlfs maklphemsp alhrkisvqy gdafhprpfc mrrvkemday
      241 gmvspfarrp svnsdlkyik akvlreggar erdkctiq

>CRA|18000005238449 /altid=gi|5059122 /def=gb|AAD38928.1|AF134409_1
          (AF134409) Rhes protein [Rattus norvegicus] /org=Rattus
          norvegicus /taxon=10116 /dataset=nraa /length=266
          Length = 266

 Score = 512 bits (1304), Expect = e-144
 Identities = 252/266 (94%), Positives = 256/266 (95%)

Query: 1   MMKTLSSGNCTLSVPAKNSYRMVVLGASRVGKSSIVSRFLNGRFEDQYTPTIEDFHRKVY 60
            MMKTLSSGNCTL+VPAKNSYRMVVLGASRVGKSSIVSRFLNGRFEDQYTPTIEDFHRKVY
Sbjct: 1   MMKTLSSGNCTLNVPAKNSYRMVVLGASRVGKSSIVSRFLNGRFEDQYTPTIEDFHRKVY 60

Query: 61   NIRGDMYQLDILDTSGNHPFPAMRRLSILTGDVFILVFSLDNRESFDEVKRLQKQILEVK 120
            NI GDMYQLDILDTSGNHPFPAMRRLSILTGDVFILVFSLD+RESFDEVKRLQKQILEVK
Sbjct: 61   NIHGDMYQLDILDTSGNHPFPAMRRLSILTGDVFILVFSLDSRESFDEVKRLQKQILEVK 120

Query: 121 SCLKNKTKEAAELPMVICGNKNDHGELCRQVPTTEAELLVSGDENSAYFEVSAKKNTNVD 180
            SCLKNKTKEAAELPMVICGNKNDH ELCRQVP EAELLVSGDEN AYFEVSAKKNTNV+
Sbjct: 121 SCLKNKTKEAAELPMVICGNKNDHSELCRQVPAMEAELLVSGDENCAYFEVSAKKNTNVN 180

FIGURE 2, page 2 of 3

```
Query: 181 EMFYVLFSMAKLPHEMSPALHRKISVQYGDAFHPRPFCMRRVKEMDAYGMVSPFARRPSV 240
            EMFYVLFSMAKLPHEMSPALH KISVQYGDAFHPRPFCMRR K   AYGMVSPFARRPSV
Sbjct: 181 EMFYVLFSMAKLPHEMSPALHHKISVQYGDAFHPRPFCMRRTKVAGAYGMVSPFARRPSV 240

Query: 241 NSDLKYIKAKVLREGQARERDKCTIQ 266
            NSDLKYIKAKVLREGQARERDKC+IQ
Sbjct: 241 NSDLKYIKAKVLREGQARERDKCSIQ 266
```

Hmmer search results (Pfam):

Model	Description	Score	E-value	N
PF00071	Ras family	126.2	2.8e-36	1
CE00060	CE00060 rab_ras_like	20.7	0.00013	2

Parsed for domains:

Model	Domain	seq-f	seq-t		hmm-f	hmm-t		score	E-value
CE00060	1/2	22	52	..	26	56	..	16.9	0.0014
CE00060	2/2	162	183	..	159	181	..	3.3	7.5
PF00071	1/1	21	186	..	1	169	[.	126.2	2.8e-36

FIGURE 2, page 3 of 3

```
   1 CTCTCTGACT CTTTGCCTCC TCTCTGACTC CCTGCCTCCT CTCTCTGTCT
  51 CCCTGCCTCC TCTGTCTGAC TCCCTGCCTC CCCTCTCTGT CTCACTGCCT
 101 CCTCTCTCTG ACTCTCTGCC TCCTCTCTCT GACTCCCTGC CTCCTCTCTC
 151 TGATTCCCTG CCTCTTTGAC CCTCTGCCTC CTCTCTTTGA CTCCCTGCCT
 201 CCTCTCTCCG ATTCTCTGCG TCTTTGACTC CCTGCCTCCT CTCTCTGACT
 251 CCCTGAAGCT CATTCAGTCA TTGCTATCAA CTCGTCTGTA CCAAGCTCTA
 301 GGCTGGAGGC TGGGCAGGGC AATGATGGAG ACAAATACTG TCCCTGGGAG
 351 CTTCTGGCCC CTTTCCCATC CTGTTTAGAC AGAAGTGACC GCCAGCAGAG
 401 TCAAGCTGTC TGCAGAAGGA CTTGGGGAGG GGGCTGTCAT GGGGTAGGGC
 451 TTCTTTCCCC CCATCTCTGC TGAAGGCCCA GGCTGGCTGA GACAGCCCCG
 501 GCAGAGACTG AGAAGGGCTC CCTGCTGTGG TCTGGCAGCC CCCTCTCCAC
 551 CCTCCTCTCT CTCATTTCCT GCCTCCCACA CGTATGCCCT GGGCACCTCA
 601 TCAGGGCTGC CCTAGGGGAG GGCCCTCCTT GGCACAGCCC CTGGGCCAGT
 651 CAGGTGGTTG AGGCTGAGGA GAGAAGGTCC CAGAGTGGGG CTTCAGGCAA
 701 ACCCAAAGAC AGAGCCCTTT GCCATTTGAT GAATGCACAG ACCCTTTATT
 751 GAGCCCCTGC TCTGTTCATG GCATGGCAGT TTTGTGGGAT AAAATTCAAAG
 801 ACAGCTTTAG GTGGGAGCTG GGTGGGGGAT GTGGGGGTCT TAGGCTTGAA
 851 CTACTACCCA GCCTCCTTTG TTAACCAAGT AGCTAGTCAC GTAGCCTTCT
 901 GAGCTCGGGG CAGACCACCT GGGATCAAAC CTCTCCTCTG CTGGTTACTG
 951 GCTGTGCAAC TGTAAGCAAG TAATTTAACC TCTCTGTGCC TCAGTTTCCT
1001 CATCTGTAAA TTGGAGAATA ACACCACCTG CTTTCTGGGG TTATGAAGGG
1051 AGAAATAGGT TAACATGTGT GCAGCACTTA GAACACTCTG GCATATTTTA
1101 GCTGCAAAAT GAATGCCAGC TATGATTATT CTATACTTA GTGCGGGGCT
1151 TGGCACACTG CATGGGCTCA AGTGGCAGCA GTTGTCGTCC TTGTGGCTCC
1201 AGGCCTGGGG TCCGCCGTGT GCTGAGCTGG CTTATTGTGC ACGTCCCTTT
1251 GTGATTCATT CATCGAAGTC ACATTAGTAG CTTAGAAGTG ACCGTAGTGG
1301 GAGCATTTAC GCCATGGAAA TTGGCAATAG GGCTTTTAAC AAAGGTATTT
1351 TTGAGAGCCG GTTTCCTGCA CAGAGGCTGG TAGTTGGGCA GGGTGAGCAG
1401 ATCCAGATGT GTGCCAGGGA CTCGCACGCA GGCAATCTCT CCACCTCCAG
1451 TGGCCATCTC AGACCTTAGC TTCATGATAG CCAGGAAGCG ATGGTGTTGG
1501 AAAGCGCCTT GGGTCAATGG GCGAGGCACT CAAGGAAACC GACTTGGGGC
1551 ATCCTGGGGT GGGGACCGAG TTTGGGCACA TACAGCCCTT TGTGTGAATT
1601 TAAAAACAGT GCCTTTTCCT CTACACAAGA TGCCCTTTCG TCTGGGATAC
1651 AGCCCCCACC TCTGGGATGC AGCCCCCACT TGCCCACCCA GCCATGCGCC
1701 TTGTGCAGTA TCCAACCTGC ACAACCTGTG GCAGCCTGTG AAGACCGAG
1751 GGGATTGATA TTTCAGCAGG CCTGTGCCCA TTTGCAGTTC AGGGGCTGGA
1801 AAGCTCTCCT CTGGAGAGGG GAGGGATTCC TGCAAGGGTG AGGAGATCAG
1851 AGAGGCCTTC AGAGAGCAGG TGGCACTTGA GCCAGACCCT GAAACATAAG
1901 GGGAAGAGGG TGTTCTGCAG AGGGGTGGCA TGAGCAAAGG AGTGGAGGCT
1951 GATCTCAGCA GAGCTCAAAC TGACGAGGGT GACTGGGGTC AGGGGTTCTG
2001 GGGCGGGGAT TCTGGTGGGC GCTAAGGTAG GAAAGGAGGG AGGGCTGGGC
2051 TGTGAAGAGC CTTTGGGGTG AGCCTGGTGG AGCCTGCGGG TTTGCTTATA
2101 CAAGAGCTTG GATCCATGTC GGCCTCTTTC ATGAGGTCAA GAGGCTCCCA
2151 TAGAAAGCTC TGAGTTTGCC CCAGAACCAT AACCCTTGGA GATGGGAGGG
2201 AAGCTTGAGC CAGCCATGGG TCGTTCCCCA TTCCACATCC TCTACTCCGG
2251 GCCTCTGGGT CTCCTGGAGG CAAGTAAACA CCTAGGGCCT GGGAGGCAAA
2301 AATATCCGGG CAGGTCATGG AGCGGAGGGA GCCCGCCAGA TGCAGAGCAC
2351 AGGTCTAAAG GTGGGTCCTC CTGAGGTGGC TGCAGGAGCA ACCCCAGGCA
2401 TTGGGCTTGG AGCATGCGGT GTGGACATAG CCTTCCCTTC TTCCCAGGAG
2451 GGCTGAATGG CCACAGAACC ACCCCCTGCC CCAGGCTTAA GAAATGCATG
2501 CTAGTGCCTT CCCCATGTCT TATCCTAGAA TCACAGGCTC CGGGAAAGCC
2551 AGATGGATGA ACCAGGGAAA GAACGGATTC TCACCATAGA TACCATTTTT
2601 GAGATTTCAC CATGTGCTGA GCCCTTTGCA ACAACTCTAT GAATTGGGCT
2651 CATTTTGCAG ATGAGAAAAG TGACTTCTAG AGAGGTTAAG CTACTAGCCC
2701 AAGATCAGTA GCTAGAGGCA AGGCAAGGAT TCAAATCCCA GGAGTCCGGT
2751 GCTTGCATAA ATGAAAGGAT GAATGAACGG ATATTGAGTG AGTGAGTGGA
2801 TGAAGGAAGG AGTAAAGGAG AGGGCATGAA TGAATGAGAG GGTAGAACTC
2851 CAAGACCCCT TAGAACCTCG TCTGATGTTC CCATTTTACA GACAGAAAAC
2901 TGAGTCCTAG ACAGAGGCCT AGAGGAGGCC AAGAGGTGGT GGGGCCAGGT
2951 CGGGGGGGCC CTGATGCCTG CTTCTCTCGC TTTGTTGCAG CCCCGAGCCA
3001 TGATGAAGAC TTTGTCCAGC GGGAACTGCA CGCTCAGTGT GCCCGCCAAA
3051 AACTCATACC GCATGGTGGT GCTGGGTGCC TCTCGGGTGG GCAAGAGCTC
3101 CATCGTGTCT CGCTTCCTCA ATGGCCGCTT TGAGGACCAG TACACACCCA
```

FIGURE 3, page 1 of 6

```
3151  CCATCGAGGA CTTCCACCGT AAGGTATACA ACATCCGCGG CGACATGTAC
3201  CAGCTCGACA TCCTGGATAC CTCTGGCAAC CACCCCTTCC CCGCCATGCG
3251  CAGGCTGTCC ATCCTCACAG GTGAGGCCCA CTGGTGCCTG GGCTGGGGCG
3301  GCAGGGCCAG GGCATGGGTG CGGAGTGTGC TGGGCACTTG GCAGTTTGCA
3351  TAGACTTGCA TAGCCATCGT CTGAGACAGG CGTCATCCCT GCACAATGAG
3401  GCTCAGAGAG GTTTTGCCAT GTGCTGGAAA TAGTGATGAA GTCGGGGGCC
3451  CCGATTCCAT TCTGTTAGAC TCCAGATCGA TTACTCATGG CTGTCGGGGC
3501  CGCCTTCCAG ATCAGGAGCT GATACCAGCA TGCCCCAGGG ATATTCCTTT
3551  CTAGGGAACA GAATGATGCC CTGGCTGCTG CTTTCCTTCT CCGGAAGATG
3601  ACCCACCAGA GCTCCAGGGC CCAAGGTCAG TCCACGGGGC TCAGGTCTCC
3651  CACACCCCAG GCCTTTGCCA CCTCCTAGAG AGGTAAGGGC AGGACCCAGG
3701  CAGTGATCAC CAAAGGGAAG GGGGCTTGGT CATGGTCATA GTGATGGTGA
3751  TGGCACTAGC TGACACTTAT CAGAAGCTAT GGGCCTGGCC CTGTTCTTAG
3801  AGCTTGGCAT GTAGTTTTTT TTGAAACAGA GTCTCGCTCT GTCACCCAGG
3851  CTGGAGTGCA GTGGCGCGAT CTCGGCTCAC TGCAACCTCT GCCTCCCGGG
3901  TTCAAGCGAT TCTCCTGCCT CAGTCCCCCA AGTAGCTGGG ACTACAGGCA
3951  CGTGCCACCA TGCCCGGCTA ATTTTTTGTA TTTTTACTAG AGACGGGATT
4001  TCACCATGTT AGCCAGGATG GTCTCGATCT CCTGACCTCG TGATCTGCCT
4051  ACCTCAGCCT CCCAAAGTGC TGAGATTACA GGCGTCAGCC ACCGCGGCCA
4101  GCCAGCATGT AGTTATTTAA CCCTCACAGT AAATAGTTAT TCATTCCCTT
4151  TTTACAGGTG GGGAAACTGA AGCCCAGAGA GGTTAAGTAA CTCACTCCAG
4201  TGGTAGCACA GCTCGTAAAG GCAGTCTGCT TTTGTGCTTT CAGACAAAGC
4251  CATACCACAG CCTCTCAGCC CTGCTGGGAA GGGTGAGGAG GGACAGGGAG
4301  GTTGGGGGGA AGAAGGGGTG AGTGGAGCTG AGGGGCTGTG CCCTTGTTTA
4351  CACTGCATTA GCATGGTAGC TAAGAGGACA AGCCCGGGAC CCAGCACCTG
4401  GGTGTGAGCC CTGGTTCCGC TGCTTCCTGG CTTTGTACCT CGAGGCAAGG
4451  GATTTTATCT CCTTGTGTCT CAGCACTCTC ATCTGTAAGA CTGCACCACA
4501  TCAACACTCA TCCTAAAGGG ACTGTGAGAC TTAAATGAAT GAATATATGT
4551  AAGGCGCTTG GTGAGCAGAT AGTAAATGCA CAATAAATCC CCAAGTCTTC
4601  TGTTGAGTCA GCATTTGCAA GTGGGCCTGC TACGGGTTAC ACGATCATTT
4651  CCCAAGTCAC GCCCCTGAAG TTGCTGAGCA GGGATAAGGG AAGGAGTGAG
4701  CAGGCAACTC TCTAGGCATC ATTCAGATAA CCCCCCAACT GAGGTACTTC
4751  TATACAGAGA AACCCATGCC ACTCCCAGCC CTGCTGCCGC CTTGAGGCCA
4801  AGACTGAGGC TGCGGGGTGG CCCCTCCTTG AGTGCTTTCT CTTCCAGGCT
4851  GGCTTTTCCG AGCATCTGAC CCAGACAGCA GTCAAGTTCT CCGCTCCACC
4901  CCGAGTTTTG GAGAAGGGGC ATGTAGATAG GAGAGCCCTG GGTCGACCTG
4951  TGTTCGAATC CTTGCTGGGC CTCTTGGCTT AAATGTGTGA CCAGAGGCAC
5001  ATGCATCCTG TCTGAGTCTC AGCCTCCCCA GCCACACAGT GGGCTTAACC
5051  TCATACCCCG CAGGGAGGCT GTGAGGACTG CAAGAAGGCT TGTGGCGGGA
5101  GCTTCCAGCA CGTGACGGGT ATTGCATTGG TGTCAGCTCC CCCAGCCTTG
5151  GGGAGGGGAC TGGGTACCCG CTGCAATGAA TAAGGCTAAT GACAGAGGGA
5201  AGGAGAGGGG AGATGTAGAG AGGAAGCACA TGCATATTTT CAGCATTAAT
5251  TTTCAGTGAC ACAAGTAATA CCCAAACACA CCCTCCTGCA AACGCTACAG
5301  ATAAAGCTAA TGCCCCTTTG ACCCATGTCC CCAATCCCAG GCTCCTGCCC
5351  CTGCCCCGGA GGTGGCCACC CTGGCAGTCT GGCATGGAGC CTTCCGGGCC
5401  TCCGTGACTA CACCGGCATT CGTATTTGTA TCCCCACAAT GGAGAGTATT
5451  TTTGTCTGTC TCTTTTTTAT GGCGCATATC ATTCTGAGCA CAGCTGTCTG
5501  ATGCTTGTTT TTTTTTCACA CACCAACCCG TGCCTCATTT TCCAACCTGG
5551  TGGAACCTCA TTTTTTCAAC CTCATTTTCC TGCTGCTCAG GAAATTCTGA
5601  AAGCCATTAA TTCCACTGCC AGCTCTTCTT CCAGCTGCCA GACGGGCCGA
5651  TCTCTGATGC TTGGCATCGC AGTCTCGCAT TTGAATATGT CAAGGCCACG
5701  AGTCCTCAGG GGCCCCGGAT TAGCCTAACG GGATGGGGTT TGGCAGCCCA
5751  AGCAGGAAGA GTTGCCAAGC TGACGCTGCC TCGCAAGTGC CTTTCAGAAG
5801  AGCCCACACT GCAGTTCCCT CCATCACCTC CCATCCATTC AGGCTTCCTT
5851  GGTTAACACT GACTGTGTGC CAGGCCCTGG GGAGACCAGG ACGAGTGGGT
5901  GATGGAACCC TTCTCTGTGC CCGAGCTGTT TGGAGCACAC CTTTGATCTG
5951  GACACCATTC TGAATGTGCC ATGTGCCATT AAATGGGGGT AAATGATGTG
6001  CTCTGGGGGT GCAGAGGAAG GTGGCAGCCA TTCTGCCAGA AGCTGGAACT
6051  GGTTGCTTCT CTTCTCAAGA ATTTGGGCCA ATTGCTGATT CCTCTGGGCC
6101  TCAGTTTCCT CATCTGTGAG ACAGGGATCT TGTCACACCA CAAGGCTATC
6151  AAGAGTTTGA GCAAAGTGG TTGGACGCAG TGGCTCATGC CTGTAATCCC
6201  AGCTCTTTGG GAGGCCGAGG TGGGCAGATC TCTTGAGGTC AGGAGTTCAA
6251  GACCAGCCTG GCTAACACAG TGAAACACCG TCTCTACTAA AAAATACAAA
```

FIGURE 3, page 2 of 6

```
6301 AAATTAGCCA GGTGTGGTGA TGGGCACCTG TAATCCCAGT TACTCGGGAG
6351 GCTGAGGCAG GAGAATCTCT TGAACCCAGG AGGTGGAGGT TGCAGTGAGC
6401 TGAGATCTTG CCATTGCATT CCAGGCTGGG CAACAAGAGT GAAACTCTGT
6451 CTCAGAAAAT AAATAAATAA ATAAATAAAA AATAGCTAGG CATGGTGACA
6501 GGCGCCTGTA ATCCCAGCTG CTCCAGAGGC TGAGGCAGGA GAATCGCTTA
6551 AACCCAGGAG GTGGAAGTTG CAGTGAGCCA AGATCACACC ACTGCACTCC
6601 AGCCTGGGCC ACAGAGAAAG ACTCCATCTC AAAAAAAAAA AAAAAAAAAA
6651 AAAAAGTTTA AGCAAAAGTG AGGAAGGTGC TTATTAAAAG CTGGAAATCA
6701 GGATGGAGGT ACCAGTCCAG ACAGCCTCCC CACCACCCCA CCGTCTCCAC
6751 AGCAGCCCCT GTTTCAGATT CACAAGCCTG CCTTGAGTGA TGCAGTGAGT
6801 TATCCTGGAG GCAGTGTGGG CCTTGGAGGC CAGCACTCAC TTTTTCATCC
6851 TATGATTTAT TTGAGAAGCA GAGAGCACCT ACCGGGTGCC AGGAACGAGC
6901 TAGGTGAGAA CAGAATCAGG TAGAAATCTC AGCCTAGCCA CACGGAAGCT
6951 GTGTGATCTT GGGCAGGCTG CATACCCTTT CTGAGCCTCA GTTTGCTCAC
7001 CTGTAATGCA AAGGTAACAA AATCTTGACA GAGGCATAGT GAGGAATCAA
7051 GAGAACAACG GGCCTGGAGC ATACACCCAG TGCTTAGCCC CCAGTAGGCC
7101 CTCACTCTCA TCATTACTGA CACCTGAGGT CACTGAGCAT GTGCCACTGT
7151 CCATTCATTA TCTTGCATAA CTCCCAAAAT CATCCTGCAA GGTAATATTT
7201 CATCTTCATG AAACAGACAG AGAAACTGAG GTTACAGAGG TTTCGTGATC
7251 TGCCCAAGTC TGCTGGCAGC TAAGCGGATG AGGCCAGATG CAAACTAGGC
7301 ATTGAGCAAG ACAGGCAGGA CCCCTGCTCT CATAGAAATG ATTTTTATTA
7351 TTATCTGAAC ACAGTCCACA CAAGTGACCT ACCCCTCTCC AGCCCTGCAA
7401 AGAAATGTGA AGTGAGTTAA CTGTATTTGA ACCAAGTGGT CCACGTGTTA
7451 GCTATGCGAC TGTGAACAGG GGCTTCAACC CCCTCAGCCT CAGTTTCCTG
7501 TCCTGGAAAA TAATCGCAGG GAGAATAATC GCAGCTACCC CGAAGAGTCG
7551 CTGTGTAGGT TAAAGCAGTT ATGCCGCATA ACTGCTTCAG GGCACCTGTG
7601 ACTCCCAGCT CTTAGGGCTG ATGTTCTGTG GCCAGAGGAG GGCAGGGGTT
7651 GCAGCTGGCC GGTGAACTCA CTACCTGGGC TCTCTCCCTG CAGGGGATGT
7701 CTTCATCCTG GTGTTCAGCC TGGATAACCG GGAGTCCTTC GATGAGGTCA
7751 AGCGCCTTCA GAAGCAGATC CTGGAGGTCA AGTCCTGCCT GAAGAACAAG
7801 ACCAAGGAGG CGGCGGAGCT GCCCATGGTC ATCTGTGGCA ACAAGAACGA
7851 CCACGGCGAG CTGTGCCGCC AGGTGCCCAC CACCGAGGCC GAGCTGCTGG
7901 TGTCGGGCGA CGAGAACTGC GCCTACTTCG AGGTGTCGGC CAAGAAGAAC
7951 ACCAACGTGG ACGAGATGTT CTACGTGCTC TTCAGCATGG CCAAGCTGCC
8001 ACACGAGATG AGCCCCGCCC TGCATCGCAA GATCTCCGTG CAGTACGGTG
8051 ACGCCTTCCA CCCCAGGCCC TTCTGCATGC GCCGCGTCAA GGAGATGGAC
8101 GCCTATGGCA TGGTCTCGCC CTTCGCCCGC CGCCCCAGCG TCAACAGTGA
8151 CCTCAAGTAC ATCAAGGCCA AGGTCCTTCG GGAAGGCCAG GCCCGTGAGA
8201 GGGACAAGTG CACCATCCAG TGAGCGAGGG ATGCTGGGGC GGGGCTTGGC
8251 CAGTGCCTTC AGGGAGGTGG CCCCAGATGC CCACTGTGCG CATCTCCCCA
8301 CCGAGGCCCC GGCAGCAGTC TTGTTCACAG ACCTTAGGCA CCAGACTGGA
8351 GGCCCCGGG CGCTGGCCTC CGCACATTCG TCTGCCTTCT CACAGCTTTC
8401 CTGAGTCCGC TTGTCCACAG CTCCTTGGTG GTTTCATCTC CTCTGTGGGA
8451 GGACACATCT CTGCAGCCTC AAGAGTTAGG CAGAGACTCA AGTTACACCT
8501 TCCTCTCCTG GGGTTGGAAG AAATGTTGAT GCCAGAGGGG TGAGGATTGC
8551 TGCGTCATAT GGAGCCTCCT GGGACAAGCC TCAGGATGAA AAGGACACAG
8601 AAGGCCAGAT GAGAAAGGTC TCCTCTCTCC TGGCATAACA CCCAGCTTGG
8651 TTTGGGTGGC AGCTGGGAGA ACTTCTCTCC CAGCCCTGCA ACTCTTACGC
8701 TCTGGTTCAG CTGCCTCTGC ACCCCCTCCC ACCCCCAGCA CACACACAAG
8751 TTGGCCCCCA GCTGCGCCTG ACATTGAGCC AGTGGACTCT GTGTCTGAAG
8801 GGGGCGTGGC CACACCTCCT AGACCACGCC CACCACTTAG ACCACGCCCA
8851 CCTCCTGACC GCGTTCCTCA GCCTCCTCTC CTAGGTCCCT CCGCCCGACA
8901 GTTGTGCTTT GTTGTGGTTG CAGCTGTTTT CGTGTCATGT ATAGTAGTAG
8951 AAATGGAAAT CATTGTACTG TAAAAGCCTA GTGACTCCCT CCTTGGCCAG
9001 GCCCTCACCC AGTTCAGATC CACGGCCTCC ACCCGGGACG CCTTCCTCCT
9051 CTGCTCCCAA ACAGGGTTTC CGTGGCCTGT TTGCAGCTAG ACATTGACCT
9101 CCGCCATTGA GCTCCACGGT TTACAGACAA TTGCACAAGC GTGGGGTGGG
9151 CAGGCCAGGA CTGCTTTTTT TTAATGCTCC CATTTCACAG AGGATACCAC
9201 CGAGACTCGG AGGGGACACG ATGAGCACCA GGCCCCACCT TTGTCCCCTA
9251 GCAAATTCAG GGTACAGCTC CACCTAGAAC CAGGCTGCCC TCTACTGTGC
9301 TCGTTCCTCA AGCATTTATT AAGCACCTAC TGGGTGCTGG GTTCACTGTG
9351 TCCTAGGAAA CCAAGAGGGT CCCCAGTCCT GGCCTCTGCC CGCCCCTGCT
9401 GCCCCACCAC CTTCTGCACA CACAGCGGTG GGGAGGCGGG GAGGAGCAGC
```

```
 9451 TGGGACCCAG AACTGAGCCT GGGAGGGATC CGACAGAAAA GCTCAGGGCG
 9501 GGTCTTCTCC TTGTGCCCGG GATTGGGCTA TGCTGGGTAC CACCATGTAC
 9551 TCAGGCATGG TGGGTTTTGA ACCCATAAAC CAAAGGCCCT TGTCATCAGC
 9601 TCTTAACAAG TATATTTTGT ATTTTAATCT CTCTAAACAT ATTGAAGTTT
 9651 TAGGGCCCTA AGGAACCTTA GTGATCTTCT ATTGGGTCTT TCTGAGGTTC
 9701 AGAGAGGGTA AGTAACTTCC TCCAGGTCAC ACAGCAAGTC TGTGGGTGGC
 9751 AGAAGCAAGC TAGCGCTGGG CATTCAGTAC ATACCACGAT GTGCTCCCTC
 9801 TCTTGATGCT TGGCCCCTGG GGCCTTCAGG GCTTTGGGAC ATCTTGTCCT
 9851 CAACCCTCTC CCTAGATCAG TCTGTGAGGG TCCCTGTAGA TATTGTGTAC
 9901 ACCATGCCCA TGTATATACA AGTACACACA GATGTACACA CAGATGTACA
 9951 CATGCTCCAG CCCCAGCTCT GCATACCTGC ACCTGCACCC CAGCCTTGGC
10001 CCCTGCCTGC GTCTGTGCTC AAAGCAGCAG CTCCAACCCT GCCTCTGTCC
10051 CCTTCCCCAC CCACTGCCTG AGCCTTCTGA GCAGACCAGG TACCTTGGCT
10101 GCACCGGTGT GTGGCCCGCT CTCACCCAGG CACAGCCCCG CCACCATGGA
10151 TCTCCGTGTA CACTATCAAT AAAAGTGGGT TTGTTACAAA GCCGTGTCCT
10201 TGCCCATGTG TATTTTTTGT ATTTCCAAGA GGAGGTGTGC CCCTTTCCAG
10251 ACCAAAGCTG GCCTTTCCCT CCCAAAATGC ACCTGCCGTG TACCCTGGCC
10301 CTGAGGGTCA GCACTGAGTC CACCTTCAAG TGTAAGTGTG GGGAGAGGGG
10351 GATAAGTCCC CCAGATGGAA GGTGATGCCC TCCTTCAGCC TGGCCCTCCT
10401 GGGTCCTCCG GGTGTGTGTA CCGAGGTGTC TGTGTCCACA AAGAAGGGGC
10451 CCCCGTGGAC CATTAGCTCC AGGAGGATCT CCGTGTCTGA GTTCTTTGTG
10501 ATTCCTGTAC AGCAGCAATT TCACCCGCAG GGGACAGTTG GCAATCTCTG
10551 GAAACCTTTT CCAAGCCTGG GGCTGGGGCT GCTACTCTCA TCTGGTGGGT
10601 GGAGGCCAGG GACACCATTC AGTATCCTCC AACGCACAGG ATGCCCCTCC
10651 ACCCCCACCC CACTGAGAAT TATCTGGCCT CAAATGCCAA GCGTGGGCAG
10701 CCTTACTTAG ACTCACCCCA GGGGCTGGGA CACGCCCCCA CCTGCGTGTG
10751 ATGGATTTGT TGGACCACAT TCTGGACGGA ACCCACAGCA TAAGCACTCC
10801 TGTGAAGTGA GACAGGATGT GGGTGAGGAT GGAAAGTGGA GGCTGAGGGA
10851 GAAGGTCTGG GCCCTGACCA ACACGGAATG TGCCCCCTGG GACTGAGAGG
10901 CTTCCCTGGG CAGAGGGAAA GGAGGAAGTC AGTGAGGTAA AATACTCCCT
10951 GTGTGTTTTA CCCAGCGAGT CTCACGCCAT CCTATCACCC AGCCCCGAGG
11001 GAAGCCCACT CATGTTCACC CCATCTGAGC ATTTAGGCTC AGAGAGCTCA
11051 ATATCTTGTC CAAGATGGCA CAGCTGGTGA AGTGGCAGAT CAGAGATTCA
11101 ACACCAGAGG CTGTCTGATT TCCGTCTGGC TGAAGAAAGA TTTTGCATCA
11151 GGGAGGTGGA AACCATCTGT GCTTTTGATC AGCAAATGCC ACCAGCAGGA
11201 TCAGGGAGCC AGGCCATAAA G
```

FEATURES:
```
Start:    3000
Exon:     3000-3270
Intron:   3271-7693
Exon:     7694-8220
Stop:     8221
```

CHROMOSOME MAP POSITION:
Chromosome 22

ALLELIC VARIANTS (SNPs):
DNA

Position	Major	Minor	Domain
3951	C	T	Intron
4127	C	T	Intron
4157	G	A	Intron
4513	C	T	Intron
6894	C	A	Intron
8409	G	C	Beyond ORF(3')
8437	T	G	Beyond ORF(3')
8579	T	C	Beyond ORF(3')
10292	A	G	Beyond ORF(3')
10792	A	G	Beyond ORF(3')

Context:

DNA
Position

3951   CACACCCCAGGCCTTTGCCACCTCCTAGAGAGGTAAGGGCAGGACCCAGGCAGTGATCAC
CAAAGGGAAGGGGGCTTGGTCATGGTCATAGTGATGGTGATGGCACTAGCTGACACTTAT
CAGAAGCTATGGGCCTGGCCCTGTTCTTAGAGCTTGGCATGTAGTTTTTTTTGAAACAGA
GTCTCGCTCTGTCACCCAGGCTGGAGTGCAGTGGCGCGATCTCGGCTCACTGCAACCTCT
GCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGTCCCCCAAGTAGCTGGGACTACAGGCA
[C,T]
GTGCCACCATGCCCGGCTAATTTTTTGTATTTTTACTAGAGACGGGATTTCACCATGTTA
GCCAGGATGGTCTCGATCTCCTGACCTCGTGATCTGCCTACCTCAGCCTCCCAAAGTGCT
GAGATTACAGGCGTCAGCCACCGCGGCCAGCCAGCATGTAGTTATTTAACCCTCACAGTA
AATAGTTATTCATTCCCTTTTTACAGGTGGGGAAACTGAAGCCCAGAGAGGTTAAGTAAC
TCACTCCAGTGGTAGCACAGCTCGTAAAGGCAGTCTGCTTTTGTGCTTTCAGACAAAGCC

4127   CAGAGTCTCGCTCTGTCACCCAGGCTGGAGTGCAGTGGCGCGATCTCGGCTCACTGCAAC
CTCTGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGTCCCCCAAGTAGCTGGGACTACA
GGCACGTGCCACCATGCCCGGCTAATTTTTTGTATTTTTACTAGAGACGGGATTTCACCA
TGTTAGCCAGGATGGTCTCGATCTCCTGACCTCGTGATCTGCCTACCTCAGCCTCCCAAA
GTGCTGAGATTACAGGCGTCAGCCACCGCGGCCAGCCAGCATGTAGTTATTTAACCCTCA
[C,T]
AGTAAATAGTTATTCATTCCCTTTTTACAGGTGGGGAAACTGAAGCCCAGAGAGGTTAAG
TAACTCACTCCAGTGGTAGCACAGCTCGTAAAGGCAGTCTGCTTTTGTGCTTTCAGACAA
AGCCATACCACAGCCTCTCAGCCCTGCTGGGAAGGGTGAGGAGGGACAGGGAGGTTGGGG
GGAAGAAGGGGTGAGTGGAGCTGAGGGGCTGTGCCCTTGTTTACACTGCATTAGCATGGT
AGCTAAGAGGACAAGCCCGGGACCCAGCACCTGGGTGTGAGCCCTGGTTCCGCTGCTTCC

4157   TGCAGTGGCGCGATCTCGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGCGATTCTCCT
GCCTCAGTCCCCCAAGTAGCTGGGACTACAGGCACGTGCCACCATGCCCGGCTAATTTTT
TGTATTTTTACTAGAGACGGGATTTCACCATGTTAGCCAGGATGGTCTCGATCTCCTGAC
CTCGTGATCTGCCTACCTCAGCCTCCCAAAGTGCTGAGATTACAGGCGTCAGCCACCGCG
GCCAGCCAGCATGTAGTTATTTAACCCTCACAGTAAATAGTTATTCATTCCCTTTTTACA
[G,A]
GTGGGGAAACTGAAGCCCAGAGAGGTTAAGTAACTCACTCCAGTGGTAGCACAGCTCGTA
AAGGCAGTCTGCTTTTGTGCTTTCAGACAAAGCCATACCACAGCCTCTCAGCCCTGCTGG
GAAGGGTGAGGAGGGACAGGGAGGTTGGGGGGAAGAAGGGGTGAGTGGAGCTGAGGGGCT
GTGCCCTTGTTTACACTGCATTAGCATGGTAGCTAAGAGGACAAGCCCGGGACCCAGCAC
CTGGGTGTGAGCCCTGGTTCCGCTGCTTCCTGGCTTTGTACCTCGAGGCAAGGGATTTTA

4513   TCGTAAAGGCAGTCTGCTTTTGTGCTTTCAGACAAAGCCATACCACAGCCTCTCAGCCCT
GCTGGGAAGGGTGAGGAGGGACAGGGAGGTTGGGGGGAAGAAGGGGTGAGTGGAGCTGAG
GGGCTGTGCCCTTGTTTACACTGCATTAGCATGGTAGCTAAGAGGACAAGCCCGGGACCC
AGCACCTGGGTGTGAGCCCTGGTTCCGCTGCTTCCTGGCTTTGTACCTCGAGGCAAGGGA
TTTTATCTCCTTGTGTCTCAGCACTCTCATCTGTAAGACTGCACCACATCAACACTCATC
[C,T]
TAAAGGGACTGTGAGACTTAAATGAATGAATATATGTAAGGCGCTTGGTGAGCAGATAGT
AAATGCACAATAAATCCCCAAGTCTTCTGTTGAGTCAGCATTTGCAAGTGGGCCTGCTAC
GGGTTACACGATCATTTCCCAAGTCACGCCCCTGAAGTTGCTGAGCAGGGATAAGGGAAG
GAGTGAGCAGGCAACTCTCTAGGCATCATTCAGATAACCCCCCAACTGAGGTACTTCTAT
ACAGAGAAACCCATGCCACTCCCAGCCCTGCTGCCGCCTTGAGGCCAAGACTGAGGCTGC

6894   GCACTCCAGCCTGGGCCACAGAGAAAGACTCCATCTCAAAAAAAAAAAAAAAAAAAAAAA
AAGTTTAAGCAAAAGTGAGGAAGGTGCTTATTAAAAGCTGGAAATCAGGATGGAGGTACC
AGTCCAGACAGCCTCCCCACCACCCACCGTCTCCACAGCAGCCCCTGTTTCAGATTCAC
AAGCCTGCCTTGAGTGATGCAGTGAGTTATCCTGGAGGCAGTGTGGGCCTTGGAGGCCAG
CACTCACTTTTTCATCCTATGATTTATTTGAGAAGCAGAGAGCACCTACCGGGTGCCAGG
[C,A]
ACGAGCTAGGTGAGAACAGAATCAGGTAGAAATCTCAGCCTAGCCACACGGAAGCTGTGT
GATCTTGGGCAGGCTGCATACCCTTTCTGAGCCTCAGTTTGCTCACCTGTAATGCAAAGG
TAACAAAATCTTGACAGAGGCATAGTGAGGAATCAAGAGAACAACGGGCCTGGAGCATAC
ACCCAGTGCTTAGCCCCCAGTAGGCCCTCACTCTCATCATTACTGACACCTGAGGTCACT
GAGCATGTGCCACTGTCCATTCATTATCTTGCATAACTCCCAAAATCATCCTGCAAGGTA

FIGURE 3, page 5 of 6

```
8409      CATGGTCTCGCCCTTCGCCCGCCGCCCCAGCGTCAACAGTGACCTCAAGTACATCAAGGC
          CAAGGTCCTTCGGGAAGGCCAGGCCCGTGAGAGGGACAAGTGCACCATCCAGTGAGCGAG
          GGATGCTGGGGCGGGGCTTGGCCAGTGCCTTCAGGGAGGTGGCCCCAGATGCCCACTGTG
          CGCATCTCCCCACCGAGGCCCCGGCAGCAGTCTTGTTCACAGACCTTAGGCACCAGACTG
          GAGGCCCCCGGGCGCTGGCCTCCGCACATTCGTCTGCCTTCTCACAGCTTTCCTGAGTCC
          [G,C]
          CTTGTCCACAGCTCCTTGGTGGTTTCATCTCCTCTGTGGGAGGACACATCTCTGCAGCCT
          CAAGAGTTAGGCAGAGACTCAAGTTACACCTTCCTCTCCTGGGGTTGGAAGAAATGTTGA
          TGCCAGAGGGGTGAGGATTGCTGCGTCATATGGAGCCTCCTGGGACAAGCCTCAGGATGA
          AAAGGACACAGAAGGCCAGATGAGAAAGGTCTCCTCTCTCCTGGCATAACACCCAGCTTG
          GTTTGGGTGGCAGCTGGGAGAACTTCTCTCCCAGCCCTGCAACTCTTACGCTCTGGTTCA

8437      AGCGTCAACAGTGACCTCAAGTACATCAAGGCCAAGGTCCTTCGGGAAGGCCAGGCCCGT
          GAGAGGGACAAGTGCACCATCCAGTGAGCGAGGGATGCTGGGGCGGGGCTTGGCCAGTGC
          CTTCAGGGAGGTGGCCCCAGATGCCCACTGTGCGCATCTCCCCACCGAGGCCCCGGCAGC
          AGTCTTGTTCACAGACCTTAGGCACCAGACTGGAGGCCCCCGGGCGCTGGCCTCCGCACA
          TTCGTCTGCCTTCTCACAGCTTTCCTGAGTCCGCTTGTCCACAGCTCCTTGGTGGTTTCA
          [T,G]
          CTCCTCTGTGGGAGGACACATCTCTGCAGCCTCAAGAGTTAGGCAGAGACTCAAGTTACA
          CCTTCCTCTCCTGGGGTTGGAAGAAATGTTGATGCCAGAGGGGTGAGGATTGCTGCGTCA
          TATGGAGCCTCCTGGGACAAGCCTCAGGATGAAAAGGACACAGAAGGCCAGATGAGAAAG
          GTCTCCTCTCTCCTGGCATAACACCCAGCTTGGTTTGGGTGGCAGCTGGGAGAACTTCTC
          TCCCAGCCCTGCAACTCTTACGCTCTGGTTCAGCTGCCTCTGCACCCCCTCCCACCCCCA

8579      GCCCACTGTGCGCATCTCCCCACCGAGGCCCCGGCAGCAGTCTTGTTCACAGACCTTAGG
          CACCAGACTGGAGGCCCCCGGGCGCTGGCCTCCGCACATTCGTCTGCCTTCTCACAGCTT
          TCCTGAGTCCGCTTGTCCACAGCTCCTTGGTGGTTTCATCTCCTCTGTGGGAGGACACAT
          CTCTGCAGCCTCAAGAGTTAGGCAGAGACTCAAGTTACACCTTCCTCTCCTGGGGTTGGA
          AGAAATGTTGATGCCAGAGGGGTGAGGATTGCTGCGTCATATGGAGCCTCCTGGGACAAG
          [T,C]
          CTCAGGATGAAAAGGACACAGAAGGCCAGATGAGAAAGGTCTCCTCTCTCCTGGCATAAC
          ACCCAGCTTGGTTTGGGTGGCAGCTGGGAGAACTTCTCTCCCAGCCCTGCAACTCTTACG
          CTCTGGTTCAGCTGCCTCTGCACCCCCTCCCACCCCCAGCACACACAAGTTGGCCCCC
          AGCTGCGCCTGACATTGAGCCAGTGGACTCTGTGTCTGAAGGGGGCGTGGCCACACCTCC
          TAGACCACGCCCACCACTTAGACCACGCCCACCTCCTGACCGCGTTCCTCAGCCTCCTCT

10292     AGCCTTGGCCCCTGCCTGCGTCTGTGCTCAAAGCAGCAGCTCCAACCCTGCCTCTGTCCC
          CTTCCCCACCCACTGCCTGAGCCTTCTGAGCAGACCAGGTACCTTGGCTGCACCGGTGTG
          TGGCCCGCTCTCACCCAGGCACAGCCCCGCCACCATGGATCTCCGTGTACACTATCAATA
          AAAGTGGGTTTGTTACAAAGCCGTGTCCTTGCCCATGTGTATTTTTTGTATTTCCAAGAG
          GAGGTGTGCCCCTTTCCAGACCAAAGCTGGCCTTTCCCTCCCAAAATGCACCTGCCGTGT
          [A,G]
          CCCTGGCCCTGAGGGTCAGCACTGAGTCCACCTTCAAGTGTAAGTGTGGGGAGAGGGGGA
          TAAGTCCCCCAGATGGAAGGTGATGCCCTCCTTCAGCCTGGCCCTCCTGGGTCCTCCGGG
          TGTGTGTACCGAGGTGTCTGTGTCCACAAAGAAGGGGCCCCGTGGACCATTAGCTCCAG
          GAGGATCTCCGTGTCTGAGTTCTTTGTGATTCCTGTACAGCAGCAATTTCACCCGCAGGG
          GACAGTTGGCAATCTCTGGAAACCTTTTCCAAGCCTGGGGCTGGGGCTGCTACTCTCATC

10792     TTCTTTGTGATTCCTGTACAGCAGCAATTTCACCCGCAGGGGACAGTTGGCAATCTCTGG
          AAACCTTTTCCAAGCCTGGGGCTGGGGCTGCTACTCTCATCTGGTGGGTGGAGGCCAGGG
          ACACCATTCAGTATCCTCCAACGCACAGGATGCCCCTCCACCCCCACCCCACTGAGAATT
          ATCTGGCCTCAAATGCCAAGCGTGGGCAGCCTTACTTAGACTCACCCCAGGGGCTGGGAC
          ACGCCCCCACCTGCGTGTGATGGATTTGTTGGACCACATTCTGGACGGAACCCACAGCAT
          [A,G]
          AGCACTCCTGTGAAGTGAGACAGGATGTGGGTGAGGATGGAAAGTGGAGGCTGAGGGAGA
          AGGTCTGGGCCCTGACCAACACGGAATGTGCCCCCTGGGACTGAGAGGCTTCCCTGGGCA
          GAGGGAAAGGAGGAAGTCAGTGAGGTAAAATACTCCCTGTGTGTTTTACCCAGCGAGTCT
          CACGCCATCCTATCACCCAGCCCCGAGGGAAGCCCACTCATGTTCACCCCATCTGAGCAT
          TTAGGCTCAGAGAGCTCAATATCTTGTCCAAGATGGCACAGCTGGTGAAGTGGCAGATCA
```

FIGURE 3, page 6 of 6